January 17, 2025

Sent via ECF

Hon. Ona T. Wang Daniel Patrick Moynihan United States Courthouse 500 Pearl Street New York, NY 10007-1312

RE: The New York Times Co. v. Microsoft Corp., et al. No. 1:23-cv-11195 New York Daily News, LP, et al. v. Microsoft Corp., et al., No. 1:24-cv-3285 The Center for Investigative Reporting, Inc. v. OpenAI, Inc., No. 1:24-cv-04872

Dear Judge Wang:

On behalf of all parties in the above-referenced actions, counsel for The New York Times Company hereby submits the below updated chart containing a summary of discovery disputes for discussion at the upcoming January 22, 2025 conference.

Given the number of issues included in the chart below, the parties have endeavored to list the issue categories in order of priority, alternating between Plaintiffs and Defendants.

Sincerely,

/s/ Ian Crosby
Ian Crosby

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 2 of 22

Key: Green represents Plaintiffs' new filing; Blue represents Plaintiffs' prior filing; Purple represents Defendants' filing

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
				I.	Disputes F	Related to Depositi	ions
<u>382</u> 430	1/13/25	403	Deposition Coordination Protocol	Microsoft & OpenAI	News Plaintiffs	The parties previously submitted letter briefs about their competing proposals for a deposition protocol. See Dkts. 261, 278. At the October 30 conference, this Court stated this dispute was premature. Oct. 30, 2024 Tr. at 8:21-23. Defendants have now submitted a letter brief regarding a revised proposal and Plaintiffs have responded with a counter-proposal.	Defendants' position: Defendants request that the Court enter a deposition protocol coordinating depositions in the New York Class Cases and News Cases. Defendants propose a simple protocol that caps the total number of hours for fact depositions at a number closely tied to the total number of hours that would otherwise be available under Rule 30 if the ten-deposition limit were applied on a per-party basis. Defendants' proposal also includes specific provisions for crossnoticing depositions, variable specified time limits for witnesses depending upon whether the deposition is a 30(b)(1) or 30(b)(6) deposition or Apex deposition, and also enables coordination of the timing of a deposition so that a witness can sit once in all of the pending cases. Finally, consistent with the Court's direction at the October status conference, Defendants' proposal does not require the Plaintiffs in the New York Cases and California Consolidated Class Action to formally coordinate their activity, but provides a method to coordinate scheduling whereby Defendants can seek an option for the same witness to sit for deposition once in all cases, which is consistent with the deposition protocol entered by Judge Illman in the California cases. The News Plaintiffs did not affirmatively file a letter brief with respect to their proposal, seemingly attempting to avoid a substantive response by Defendants. The News Plaintiffs' proposal would not achieve meaningful coordination and would lead to multiple depositions of the same fact witnesses on the same issues. Plaintiffs present multiple convoluted proposals that do not facilitate efficiency. In addition, Plaintiffs provide no justification for the number of hours of depositions they are seeking. For example, the hours caps sought by the News Plaintiffs alone (not accounting for the Class Plaintiffs) amount to the equivalent of nearly 40 depositions. The Class Plaintiffs separately seek the equivalent of over 30 depositions themselves. Plaintiffs have not and cannot justify departing fro

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 3 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
		No.					alternative, if the Court wishes to order coordination across all SDNY Cases, including the News cases and the <i>Authors Guild</i> class cases, the Court should enter the News Plaintiffs' alternative proposal, which is attached as Exhibit 2 to the brief. Dkt. 403-2. The News Plaintiffs' proposal is superior to Defendants' proposal in four key ways. <i>First</i> , Defendants' proposal is vague as to what coordination (if any) is required with the <i>Tremblay</i> class plaintiffs in the Northern District of California. This Court already denied Defendants' request for coordination with the California plaintiffs, stating at the October 30, 2024 status conference that it is "not going to require anybody to formally coordinate with the actions in the Northern District of California." October 30, 2024 Tr. at 6:25-7:2. The News Plaintiffs do not oppose attempting to coordinate scheduling with the California action, but Defendants' new proposal arguably seeks more, requiring an "option for the same witness to sit for deposition once in all [i.e., the New York and California] cases." Dkt. 382 at 2. <i>Second</i> , the News Plaintiffs reasonably propose a higher total hour allotment (260 versus 140), which accounts for the 65 current document custodians and long list of issues to explore during depositions. <i>See</i> Dkt. 261 at 4 n.8. <i>Third</i> , The News Plaintiffs propose a more reasonable cap of 10 hours per witness, to be shared by the News Plaintiffs only, equating to 3.33 hours for each News Plaintiff. By contrast, Defendants' proposed cap equates to an unreasonably low 90 minutes for each of the three News Plaintiffs. <i>Fourth</i> , the Court should reject Defendants' request to limit "apex" depositions to five hours. Defendants have not even said who they believe the "apex" witnesses are. Moreover, discovery to date
							demonstrates that high-level executives at OpenAI and Microsoft were deeply involved in issues that are central to this case. <i>See, e.g.</i> , Dkts. 403-5, 403-6, 403-7 (examples of documents produced by Defendants).
386	1/13/25	412	Microsoft's use of the "AEO" confidentiality designation	News Plaintiffs	Microsoft	News Plaintiffs and Microsoft have submitted letter briefs, on which the Court has not	Plaintiffs' position: In light of Microsoft's over-use of the "AEO" confidentiality designation, the News Plaintiffs seek a minor modification to the Protective Order that would permit the News Plaintiffs' to use Microsoft's AEO documents during depositions of OpenAI employees, provided such persons have first executed the agreed Non-Disclosure Agreement.

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 4 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
						yet heard argument.	As of this filing, 60.26% of Microsoft's production remains AEO. By contrast, The Times has designated 1.1% of its documents as AEO. Microsoft's designation remains closer than not to the 70% threshold that "support[s] an inference of bad faith" and justifies a full-fledged rereview of all AEO documents. <i>United States v. Planned Parenthood Fed'n of Am.</i> , 2022 WL 19010531, at *3 (N.D. Tex. Dec. 15, 2022). As a result, the News Plaintiffs are effectively precluded from using the majority of Microsoft's documents during depositions of OpenAI employees. The Protective Order does not allow parties to show AEO documents to a deponent unless that deponent is "the author or recipient of a document containing the information or a custodian or other person who otherwise possessed or knew the information." Dkt. 127 at 8. This is a problem. Given the extensive collaboration between Microsoft and OpenAI to develop and commercialize the models and products at issue in this case, many Microsoft documents discuss OpenAI employees and/or Microsoft's collaboration with OpenAI on relevant issues. The News Plaintiffs appropriately seek to use these documents during depositions of OpenAI employees. <i>See, e.g.</i> , Dkts. 386-2, 386-3, 386-4, 386-5 (examples of Microsoft "AEO" documents the News Plaintiffs may seek to use during depositions of OpenAI employees).
							Microsoft's position: This is not a dispute about over-designation of documents—the designation issues were previously resolved by Microsoft's re-review and re-production of numerous documents, among other measures, resulting in a substantial decrease in the percentage of its documents designated AEO. Additionally, Microsoft agreed that certain individuals at the News Plaintiffs could view such documents and that the News Plaintiffs could show any such documents to any Microsoft employee witness at deposition. What Microsoft cannot agree to is allowing the News Plaintiffs to show any employees of its direct competitor OpenAI any document produced by Microsoft and designated as AEO—that is a serious, and unwarranted, invasion of Microsoft's ability to protect its most sensitive business information. To be clear, what the News Plaintiffs describe as a "minor modification" of the Protective Order would essentially allow them to automatically re-designate any Microsoft AEO document to CONFIDENTIAL by presenting it to an OpenAI witness in a deposition. Microsoft has a legitimate need to protect its sensitive business and technical information from disclosure to OpenAI. The

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							News Plaintiffs completely disregard this fact and in doing so they fail to show good cause to rewrite the fundamental terms of the Protective Order.
				II. Dispu	utes Related to P	laintiffs' Requests	s for Inspection
305 ¹ 328	11/1/24 11/20/24	305 345	OpenAI's Training Data	News Plaintiffs	OpenAI	The Court heard argument on December 3, 2024, and ordered the Parties to meet and confer on the issues raised in these motions and inform the Court of the status of these issues in this joint chart. (Dkt. 365)	Plaintiff's position: The Times and the Daily News Plaintiffs have continued their training data inspection, and have asked OpenAI to run two types of searches over OpenAI's training datasets: (1) a keyword search over URL data contained in specific training datasets ("URL Search"), and (2) an n-gram search over text data contained in OpenAI's training datasets ("N-Gram Search"). For (1), OpenAI agreed to run the URL Search but as of today has not yet produced any search results to News Plaintiffs. *After* News Plaintiffs sent their position statement, OpenAI finally committed to producing the results by January 31. For (2), on November 13, 2024—more than two months ago—News Plaintiffs shared with OpenAI their specific request for the N-Gram Search (including the precise code for OpenAI to run). OpenAI has refused to run the N-Gram search, purportedly because it is concerned about the possibility of false positive hits. Based on this professed concern, on December 18, 2024, News Plaintiffs asked OpenAI to provide a counterproposal for the N-Gram Search for News Plaintiffs' review. OpenAI did not provide any counter-proposal until January 17, 2025, *after* News Plaintiffs sent their position statement. News Plaintiffs are considering that proposal and will continue to meet and confer with OpenAI. In addition, OpenAI has now committed to say whether it agrees to run the N-Gram Search by January 27, 2025. OpenAI's position: On (1), Plaintiffs' request for assistance with URL searches, there is no dispute. Plaintiffs complain that OpenAI did not commit to a production until they served their statement; but had they reached out directly, they would have learned that the searches were complete and headed to OpenAI's ESI vendor. Production by January 31 should be feasible.

¹ Unless otherwise indicated, all docket numbers are from *New York Times Company v. Microsoft Corp., et al.*, Case No. 1:23-cv-11195-SHS.

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 6 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							On (2), the ngram search, Plaintiffs' statement ignores the extensive work that has taken place. As discussed at the last hearing, the "precise code" that Plaintiffs sent was unusable because it had a memory management issue that made it impossible to run at scale on OpenAI's systems (or likely any other scalable system for distributed computing). 12/3/24 Tr. at 15-16. OpenAI pledged to have its experts "essentially prototyp[e]" an "alternative way" to write the program to see if it would even be possible to run on OpenAI's systems. <i>Id.</i> at 16. OpenAI did that in December and reported to Plaintiffs that it could fix the memory issue by rewriting the code. But OpenAI still needed two pieces of information to estimate the cost and burden of running Plaintiffs' search for them. First, Plaintiffs promised to provide a list of the datasets that it wanted OpenAI to search back in December. Just over a week ago, and once again today, Plaintiffs indicated that they would identify the datasets by name; information that is needed to assess the burden of these searches. Second, Plaintiffs failed to provide one of the key terms for their search algorithm. Right before the holidays, Plaintiffs instead asked OpenAI to make its own proposal for that term. But to make such a proposal, OpenAI has had to run a series of experiments. The experiments are now complete, the parties met and conferred to discuss, and they are continuing to meet and confer. At the last hearing, Plaintiffs asked for a date certain by which OpenAI will say whether it agrees to run this ngram search. <i>Id.</i> at 10-11. As long as Plaintiffs send the necessary dataset information, OpenAI can commit to provide that decision by January 27.
<u>379</u>	1/13/25	420	OpenAI's Log Output Data	News Plaintiffs	OpenAI	News Plaintiffs and OpenAI have submitted letter briefs, on which the Court has not yet heard argument.	Plaintiff's Position: The Times and Daily News Plaintiffs request an order compelling OpenAI to preserve output log data on a going-forward basis for OpenAI's generative AI products at issue, including ChatGPT, SearchGPT, APIs, Browse, and certain Custom "GPTs," and to identify the nature and extent of output log data which it has destroyed at least since December 27, 2023, when The Times filed its initial Complaint. It was not until OpenAI sent a letter to News Plaintiffs on November 15, 2024, that News Plaintiffs became aware that OpenAI was not preserving all relevant output log data that had long been requested. OpenAI has not explained why it cannot preserve the data—as it has a duty to do—or detailed the alleged burden of preserving the data.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							OpenAI's position: The News Plaintiffs seek an order requiring OpenAI to preserve all ChatGPT user conversation data, even if the conversation has no relationship to Plaintiffs' content, and even if a user has requested the conversation's deletion. This is an extraordinary request. Between when The Times filed their original complaint on December 27, 2023, and now, hundreds of millions of ChatGPT user accounts have been created, which send one billion messages each day. Consistent with OpenAI's Privacy Policy, OpenAI retains ChatGPT user conversation data by default, while also respecting user deletion requests and requests to opt out of having their ChatGPT conversation data retained. For purposes of this litigation, OpenAI has further retained data associated with specific ChatGPT user accounts likely to have responsive information. In light of this extraordinary volume of data, counsel have for months engaged in productive and ongoing conversations regarding how to sample and use it. Apparently not satisfied with an enormous sample of already-preserved conversations for litigation analysis, Plaintiffs now request OpenAI "preserve [all] its output log data" going forward, for all users and all conversations worldwide. See ECF 379 at 1, 3. As OpenAI explained in detail in its opposition to Plaintiffs' motion, the Court should reject this request to preserve all user conversations as (1) non-responsive, (2) unduly burdensome, and (3) untimely, given that OpenAI made The Times explicitly aware of its retention policies in February 2024 and The Times failed to raise the issue until November. Such a sweeping change to OpenAI's ordinary operations and user privacy is disproportionate to the needs of the case. Accordingly, this Court should deny Plaintiffs' request.
<u>311</u>	11/18/24	<u>338</u>	Microsoft's Log Output Data	News Plaintiffs	Microsoft	The Court heard argument on December 3, 2024, and ordered the Parties to meet and confer on processes to enable Plaintiffs to	Plaintiff's position: The Times and Daily News Plaintiffs sent a letter to Microsoft on December 18, 2024, requesting an informal discussion with Microsoft's experts regarding the log dataset, the meaning of different fields, and how the dataset can be queried. The letter outlined a list of information to be provided before the expert discussion, including questions about how the log dataset(s) are stored and documents Microsoft has produced thus far, and a set of questions to be addressed by the technical experts. News Plaintiffs did not receive a substantive response from Microsoft until January 16, 2025. The News Plaintiffs will continue to meet and confer with Microsoft in order to convene an informal discussion among the technical witnesses.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
						query the dataset efficiently in advance (and possibly instead) of using a deposition; and inform the Court of the status of these issues in this joint chart. (Dkt. 365)	Microsoft's position: Plaintiffs sent their letter on December 18, 2024, just before the holidays. In response, Microsoft performed an in-depth further investigation to enable it to respond to the questions posed in the News Plaintiffs' letter. On January 16, 2025, Microsoft provided a substantive response to the News Plaintiffs' preliminary questions, produced additional documentation explaining the data logs, and produced additional data log samples to aid in the News Plaintiffs' understanding of the data. Microsoft has offered to meet and confer with the News Plaintiffs and will continue working in good faith to answer their questions, including to convene an informal discussion with its technical witnesses as soon as practicable.
<u>250</u>	9/27/24	<u>250</u>	OpenAI's Model Inspection Protocol	News Plaintiffs	OpenAI	The Court heard argument on December 3, 2024, and ordered Defendants to provide Plaintiffs one week of "free" searches. (Dkt. 365)	Plaintiff's position: The Times and Daily News Plaintiffs conducted their week of "free" searches of OpenAI's models in December. News Plaintiffs will meet and confer with Defendants if they need additional searches. OpenAI's position: OpenAI will meet and confer with Plaintiffs if they determine they require further inspections.
				III. Dispute	es Related to Cus	stodians and Cust	odial Documents
393 – Sealed 395 – Public	1/13/25	<u>408</u>	OpenAI's request that The Times add six custodians and expand the date range for certain search terms back to 2018	OpenAI	The Times	OpenAI and The Times previously submitted letter briefing related to a portion of this dispute (Issue 1). Dkts. 316,	OpenAI's position: OpenAI seeks an order compelling the Times to: (1) add David Rubin, Michael Greenspon, Hannah Masuga, and Sam Dolnick as custodians (renewing its previous motion, ECF 316); (2) add Justin Heideman, Staff Software Engineer, and Elijah Soria, Senior Software Engineer as custodians for search terms related to the Times's paywall and other access controls; and (3) collect and search custodial files back to 2018 for RFPs related to OpenAI's alleged misconduct and The Times's allegations of harm.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
						317, 330. The Court directed the parties to meet-and-confer regarding this issue and provide an update at this status conference. Dkt. 365. OpenAI and The Times have now submitted additional letter briefing concerning Issue 1 and on the two new issues in OpenAI's motion.	First, the Times has failed to make custodial productions necessary for the parties to meet and confer—the Times has only produced 355 documents since the last hearing, 292 of which are non-custodial documents. The Times's delay of its productions is prejudicing OpenAI, who cannot afford to wait until the end of the discovery period to renew its request. Second, the Times has itself identified Heideman and Soria as knowledgeable regarding the Times's access controls in Interrogatory responses, yet for some reason refuses to designate these technical witnesses who are most likely to have day-to-day conversations about these subjects as custodians for the limited search terms OpenAI has proposed. Third, the Times misrepresents OpenAI's date range request, which it has made for targeted search terms and for specific custodians, for which negotiations have been ongoing. The Times is the only News Plaintiffs which has taken such a narrow view of the relevant date range, and has offered no evidence that it would be unduly burdensome for it to undertake this collection. Plaintiff's position: For the previously sought custodians (Issue 1), OpenAI's motion should be denied because The Times has designated numerous custodians with documents relevant to the discovery OpenAI claims to need (and from whom The Times is still producing custodial documents). For the newly sought custodians (Issue 2), OpenAI's motion should be denied based on the existing custodians and data to which OpenAI already has access, including non-custodial data that is a better source of the information they seek than the emails of two new custodians. Finally, for the date range dispute (Issue 3), OpenAI improperly seeks to re-visit negotiations that occurred last July, when The Times represented that it would collect documents beginning in 2022 (the year ChatGPT was released) and Defendants did not object. In any event, The Times has agreed to produce documents going back to 2018 for several of its high-ranking employees, Dkt. 393 at 3 n.7, and has

_

² Those custodians include A.G. Sulzberger, Meredith Levien, Alex Hardiman, William Bardeen, Zach Seward, Ben Frumin, David Gurian-Peck, Barrett Sheridan, Leilani Han, Hannah Yang, and David Perpich.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
<u>228</u>	8/27/24	233	OpenAI's search terms (The News Plaintiffs request that OpenAI run additional search terms for its custodians' ESI)	News Plaintiffs	OpenAI	Addressed at the October 30th conference and the December 2nd conference, but not yet entirely resolved	Plaintiff's position: This Court granted the News Plaintiffs' request to impose a target hit count for OpenAI's custodial document review, selecting a 600,000 target for the current set of 34 custodians. Dkt. 365. However, as described by the Court's December 6, 2024 Order, the parties still have a dispute about whether that target "applies to all custodial documents produced to date or only those produced after the October 30, 2024 conference." Dkt. 365 at 4; see also Dec. 3, 2024 Tr. at 61:3-64:21. The News Plaintiffs' position is that the Order requires OpenAI to review 600,000 additional documents, on top of the approximately 130,000 documents that OpenAI had already reviewed before The Times filed its search term motion to compel in August. In briefing that motion, both sides explicitly described the dispute as pertaining to the "additional" documents that OpenAI should have to review, as opposed to a total number of documents. Dkt. 228-1 at 1; see also Dkt. 233-2 (OpenAI declarant acknowledging that The Times's motion sought to have OpenAI review "new documents (i.e., not previously reviewed by OpenAI")). Accordingly, under the News Plaintiffs' interpretation of the Court's Order, the total hit count target for OpenAI is 730,000 total documents (130,000 plus 600,000). The News Plaintiffs have a proposal to resolve this dispute. The News Plaintiffs' most recent search term proposal to OpenAI yielded 666,620 documents, which essentially splits the difference between the parties' competing interpretation of the Court's Order. The News Plaintiffs' proposal is that OpenAI run these search terms, and that the parties going forward not request any additional search terms absent good cause. This proposal is intended to resolve all disputes about search terms once and for all, as The Times and Daily News Plaintiffs' have for many months been extensively discussing search terms with OpenAI for outgoing and incoming document productions. OpenAI understands that this cap applies to the total, all-in number of docum

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 11 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							(listing " <i>Total</i> Documents Hit by Term Including Families" (emphasis added)). OpenAI offered to meet and confer on this issue multiple times since the last hearing.
							OpenAI has been actively considering the News Plaintiffs' proposal to resolve the dispute as to <i>OpenAI's</i> custodial documents only. Plaintiffs bootstrap their proposal to "resolve all disputes about search terms," including search terms for the News Plaintiffs' documents. But the parties are not similarly situated. For example, the Center for Investigative Reporting has produced less than 400 documents to date. And OpenAI is still in the early stages of conferring with the Daily News Plaintiffs and Center for Investigative Reporting on those Plaintiffs' search terms and custodians. To require that no additional search terms can be requested absent good cause for Plaintiffs' custodial documents at this stage of negotiations would make no sense. OpenAI would be inclined to accept a version of the proposal if (1) it resolves the dispute as to <i>OpenAI's</i> custodial documents, while allowing ongoing negotiations regarding the New Plaintiffs' custodians and custodial documents to proceed; and (2) "good cause" means that Plaintiffs may only propose additional custodial search terms based on new requests for production of documents (i) served after the date of this filing, (ii) that are not encompassed within existing document requests, and (iii) that seek information about relevant issues that Plaintiffs could not have reasonably known about from information available to them at the time of this filing.
<u>381</u>	1/13/25	411	Microsoft's search terms (The News Plaintiffs request that Microsoft run additional search terms for its custodians' ESI, and that Microsoft extend the date range for collections to July 1, 2024)	News Plaintiffs	Microsoft	News Plaintiffs and Microsoft have submitted letter briefs, on which the Court has not yet heard argument.	Plaintiffs' position: The News Plaintiffs seek an order (1) imposing a 600,000 target hit count for Microsoft's custodial ESI review, and (2) requiring Microsoft to collect documents through July 1, 2024. In its opposition brief, Microsoft agreed to the July 1, 2024 date range, mooting that portion of the dispute. Dkt. 411 at 3. But there is still a dispute about issue 1—the hit count target. After six months of negotiating search terms, it is time for these negotiations to conclude so the parties can move on to the next phase of the case. The best path forward is a familiar one: this Court should resolve this dispute the same way it resolved the News' Plaintiffs' analogous dispute with OpenAI—that is, to impose a target hit count and direct the News Plaintiffs to tailor their proposals to meet that target. See Dkt.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							365. This approach promises to break the parties' current logjam, as Microsoft has repeatedly rejected the News Plaintiffs' proposed terms without making compromise offers.
							In its opposition brief, Dkt. 411, Microsoft has provided additional information about its document review efforts for the first time, which the News Plaintiffs will meet and confer about with Microsoft in advance of the January 22 hearing.
							Microsoft's position: Microsoft's central focus has consistently been to reach a final, mutually agreeable set of search terms that is appropriately narrowed to capture documents relevant to this case, rather than an agreement as to an arbitrary number of documents for review. To achieve these ends, Microsoft has willingly and cooperatively engaged in good faith negotiations with News Plaintiffs. However, the News Plaintiffs have continued to propose search terms that are far too broad and encompass issues outside of this case resulting in over a million parent-level hits, notwithstanding the requested inclusion of 6 months of additional documents and the additional search terms News Plaintiffs intend to add to address some of the other discovery issues before the Court.
							It is not clear to Microsoft what relief News Plaintiffs now seek given that Microsoft has already reviewed over 600,000 custodial documents for production, which has resulted in production of more than 132,000 documents consisting of more than 1 million pages, with additional custodial documents in the collection and production queue. It may be that News Plaintiffs are requesting review of an additional 600,000 documents hitting only their own search terms. This would effectively require Microsoft to review over 1.2 million documents given that the hit counts Microsoft provides in response to the search term requests are unique and de-duplicated against documents that have already been collected for review.
					10		Microsoft maintains, as it has proposed to News Plaintiffs, that the parties should continue to focus on the substantive aspects of custodial search term negotiations as opposed to an arbitrary hit count cap that has already been met by Microsoft. As to News Plaintiffs' proposed date range cutoff, Microsoft agrees to run News Plaintiffs' requested search terms on documents up to and including July 1, 2024, based on

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							News Plaintiffs' agreement that these documents will count toward any cap imposed by the Court or agreed upon by the parties.
204 313	8/12/24 11/18/24	211 342	OpenAI's custodians (The News Plaintiffs seek additional OpenAI custodians)	NYT	OpenAI	Addressed at the October 30 conference and the December 2 conference, and resolved in part	Plaintiffs' Update: This dispute has been resolved in part, but certain disputes remain. Prior to the December 2 conference, the News Plaintiffs were requesting 12 additional OpenAI custodians. See Dkt. 204 (requesting 8 additional custodians); Dkt. 313 (requesting 4 additional custodians, in addition to the previously requested 8 custodians). At the December 2 conference, this Court granted the News Plaintiffs' request for the 8 custodians addressed in Dkt. 204. See Dkt. 365. As to the 4 remaining custodians, this Court granted the News Plaintiffs' request in part, requiring OpenAI to cross-produce documents from the Tremblay case in N.D. Cal. for two of the four custodians (Michael Lampe and Suchir Balaji), and directing the parties to meet and confer once those productions were complete to discuss (1) whether additional productions for Lampe and/or Balaji were warranted, and (2) the remaining two disputed custodians (Peter Deng and Ethan Stock). Dkt. 365; Dec. 2, 2024 Tr. at 53:1-59:19. OpenAI has not yet completed the court-ordered productions for the new custodians. Once completed, the News Plaintiffs will confer with OpenAI to discuss any need additional productions for the four custodians who remain disputed: Lampe, Balaji, Deng, and Stock. OpenAI's position: Since the last agenda update, OpenAI has produced 16,884 documents. The Times, in contrast, has produced just 355 documents since then. Production remains in progress because, among other things, the parties are still negotiating search terms. OpenAI will meet and confer with Plaintiffs about whether additional discovery is warranted once the productions are complete and Plaintiffs have had a chance to review them to determine what, if any, additional relevant, responsive, non-duplicative documents they require.
309	11/18/24	329- OpenAI 331- MSFT	Defendants' preservation of, search for, and production of	News Plaintiffs	OpenAI & Microsoft	News Plaintiffs and Defendants submitted	Plaintiffs' Update: This dispute has been resolved in part, but one dispute remains. This Court granted the News Plaintiffs' motion to compel Defendants to produce text and social media messages. Dkt. 360. That ruling,

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			text messages and social media messages			letter briefs in advance of the December 3, 2024 conference, and the Court resolved the motion in part (Dkt. 360)	however, did not address the News Plaintiffs' dispute with Microsoft concerning the format in which text messages should be produced. The News Plaintiffs have requested that productions be made in 24-hour increments, while Microsoft has taken the position that productions need only be made in a "usable format for plaintiffs." Dkt. 360 at 2. At the December 2 conference, the News Plaintiffs agreed to table this format dispute until after Microsoft produced text messages. Microsoft has not yet produced any text messages, so this dispute has not yet been resolved.
							Microsoft's position: Microsoft disagrees that there is a live dispute for the Court's consideration, as this was addressed at the previous conference. Microsoft is still in the process of collecting and reviewing text messages, and will produce any responsive documents in due course.
				IV. Disp	utes Related to C	OpenAI's Docume	nt Productions
398	1/13/25	419	The Times's RFPs 113-14 (documents concerning the market for model training data and the market for content for retrieval-augmented generation ("RAG") or generative search; The Time's RFPs 67, 70, 99, 100 (documents concerning OpenAI's revenues and profits, including financial presentations,	NYT	OpenAI	The Times and OpenAI have submitted letter briefs, on which the Court has not yet heard argument.	Plaintiff's position: This motion to compel addresses three categories of RFP disputes with OpenAI. First, OpenAI should be required to search for documents that explicitly address the "market" for model training data and RAG data. Despite agreeing that such documents are relevant to fair use, OpenAI initially refused to conduct this search, arguing these documents will be privileged because only lawyers use the word "market," and that any review will be burdensome. Dkt. 398-3 at 4. OpenAI has now reconsidered that position, as reflected in its opposition brief, but the parties still have a dispute about which search terms should be used to collect responsive documents. Second, OpenAI should produce documents concerning revenues and profits for the at-issue models and products, which are relevant to motive, OpenAI's commercial purpose, and damages. Specifically, OpenAI should be ordered to produce financial presentations and related communications regarding its revenues, profits, and valuation, as well as about Microsoft's investments in OpenAI. Such documents will provide relevant commentary and analysis lacking from the financial statements that OpenAI has agreed to produce. OpenAI should also be ordered to produce information about its projected revenue streams and future profits, which are recoverable. See Looney Ricks

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			projected revenues, and referral revenue); The Times's RFP 106 (documents OpenAI has submitted to domestic or foreign entities relating to the allegations in the Complaint)				Kiss v. Bryan, 2010 WL 5175167, at *2 (W.D. La. Dec. 7, 2010) ("[T]his Court finds as a matter of law that copyright holders can pursue future profits"). OpenAI should also be ordered to produce information about its referral, affiliate, advertising, and paid search revenue, which are also relevant to The Times's damages. Third, and finally, OpenAI should identify any domestic or foreign governmental authorities to whom OpenAI has produced documents that are relevant to the allegations and defenses in this case, and OpenAI should produce those documents. OpenAI has refused to even investigate the existence of a non-U.S. proceeding in which it has produced relevant documents, despite public information suggesting there may be such proceedings. See https://www.theguardian.com/technology/2024/jan/08/ai-tools-chatgpt-copyrighted-material-openai (discussing a submission to the House of Lords in which "OpenAI said it could not train large language models such as its GPT-4 model – the technology behind ChatGPT – without access to copyrighted work."). OpenAI's position: The issues raised by this motion to compel are premature and/or inappropriate. First, OpenAI agreed to the sole search term referenced in The Times' motion and the only search term identified as related to RFPs 113/114. OpenAI understood that this would resolve the dispute. Two days after filing its motion, The Times moved the ball and proposed eight additional terms never previously identified as related to these RFPs. OpenAI continues to consider The Times' proposal, but objects to The Times' attempt to shoehorn additional disputes into this motion. Second, The Times' request for more documents concerning OpenAI's revenues and profits is a fishing expedition—The Times already has OpenAI's financial statements which contain extensive notes on how OpenAI's financial figures were calculated. The Times fails to explain why further evidence is relevant and proportional to the needs of the case, or why it is non-duplicative. Judge Illman denied a substa

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							(N.D. Cal. Sept. 30, 2016). The Times misleadingly cites <i>Looney Ricks</i> as support for its position, but that case is inapposite and contrary to the rulings of courts in this Circuit. To the extent that The Times requests information about OpenAI's "referral, affiliate, advertising, and paid search revenue," OpenAI has already offered to produce evidence responsive to this request. But the Times again asks for more without explaining why existing productions are insufficient or why its requests are not duplicative. Third, The Times' only justification for documents submitted to any government worldwide is that such evidence is relevant to subjects like OpenAI's model training practices. It strains credulity for The Times to argue such an overbroad request is proportional to the needs of the case, especially because The Times has served 130 RFPs, including 20 that address model training.
322- Sealed 323- Public	11/19/24	339	The Times's RFP 49 (documents concerning Defendants' knowledge of whether Defendants' models contain copyrighted content); The Times's RFP 2 (documents OpenAI has submitted to domestic governmental entities relating to the allegations in the Complaint)	NYT	OpenAI	The Times and OpenAI submitted letter briefs in advance of the December 2 conference; the Court did not hear argument on this dispute at that conference.	Plaintiff's position: For RFP 49, The Times requests an order requiring OpenAI to produce documents "concerning Defendants' monitoring, tracking, and knowledge of investigations into whether Defendants' AI Models contain copyrighted content, including Times Content." This request is relevant to The Times's allegation of willful infringement. FAC ¶¶ 124-26. The dispute has been narrowed to whether OpenAI should limit productions to documents that specifically address The Times, the Daily News Plaintiffs, and the CIR. The Times appropriately seeks a broader scope, which should include documents addressing OpenAI's awareness that it was using copyrighted journalism content more generally. OpenAI has now reconsidered its position, and the parties are very close to resolving this dispute. The Times awaits OpenAI's response to its latest proposal, which addresses the definition of "journalism" that should apply to this RFP. For RFP 2, The Times requests an order requiring OpenAI to identify any relevant proceedings in which it has produced documents related to the issues in this case, which will facilitate additional negotiations about the appropriate scope of OpenAI's response to RFP 2. OpenAI has contended "there are no other U.S. government proceedings where OpenAI has produced documents that concern or relate to the allegations in the Complaint." But The Times cannot accept that representation because OpenAI has clarified it is excluding from its response highly relevant issues, including for example investigations about OpenAI's transition to a for-profit company. See

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 17 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							https://www.axios.com/2024/10/30/openai-for-profit-delaware-attorney-general; A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1015 (9th Cir. 2001) ("This 'purpose and character' element also requires the district court to determine whether the allegedly infringing use is commercial or noncommercial.").
							OpenAI's position: For RFP 49, OpenAI confirmed in the briefing on this issue that it is <u>not</u> limiting its productions to documents that specifically address The Times, the Daily News Plaintiffs. ECF No. 339 ("If The Times were to define journalism as the 'activity of writing for newspapers, magazines, news websites or any other [written] publication,' then OpenAI would agree to that scope."). Regardless, OpenAI agrees with The Times that the parties are likely to resolve this dispute and that it does not require intervention from the Court. With respect to RFP 2, there is nothing for the Court to decide. As OpenAI has previously confirmed, there are no relevant U.S. government proceedings where OpenAI has produced documents. The
							Times's citation to a news article about a reported inquiry by a state regulator concerning OpenAI's corporate structure has no relevance to this case. The Times suggests that this reported inquiry is relevant to OpenAI's "transition to a for-profit company," but OpenAI has had a for-profit entity since 2019. Therefore, any regulatory inquiries into changes to its corporate structure now are wholly irrelevant. Moreover, the Times's request for an "order requiring OpenAI to identify any relevant proceedings" is improper. An RFP is a request for a production of documents; it is not a basis to demand a narrative response, much less a response concerning irrelevant issues.
				V. Dispu	ites Related to M	licrosoft's Docum	ent Productions
<u>396</u>	1/13/25	<u>407</u>	The Times's RFPs 113-14 (documents concerning the market for model training data and the market for content for retrieval- augmented	NYT	Microsoft	The Times and Microsoft have submitted letter briefs, on which the Court has not yet heard argument.	Plaintiff's position: This motion to compel addresses three categories of RFP disputes with Microsoft. First, The Times seeks an order requiring Microsoft to produce internal and external communications concerning Microsoft's negotiations about content licensing agreements, and documents about the existence of a market for the at-issue training and RAG data. Microsoft concedes such documents are relevant to fair use, and has agreed to produce some of the requested documents. But Microsoft has improperly refused to produce (1) documents concerning negotiations that never resulted in a

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			generation ("RAG") or generative search; The Times's RFP 123 (documents concerning Microsoft's investments in OpenAI); The Times's RFP 109 (documents concerning the impact of Defendants' products on copyright owners).				final agreement, and (2) documents that address the market for the atissue data without focusing on a specific negotiation with a specific licensing partner. In its opposition brief, Dkt. 407, Microsoft has reconsidered its position on this second issue. However, the first issue (communications regarding potential licensing agreements) remains disputed, and Microsoft's reliance on the <i>Authors Guild</i> ruling is misplaced for the reasons explained in The Times's brief. Second, The Times seeks an order requiring Microsoft to produce documents about its investments in OpenAI, which are relevant to The Times's contributory and vicarious liability claims against Microsoft. FAC ¶ 66, 151, 169. Microsoft has agreed to produce some documents in response to this request, but has objected to providing discovery about its 2024 investments on the ground that The Times's case was filed in December 2023. But the News Plaintiffs allege ongoing misconduct, requiring discovery into the ongoing commercial exploitation of their content. See, e.g., Times FAC ¶ 167. Moreover, even if it were appropriate to limit discovery as of the Complaint's filing date, the consolidated Daily News and CIR cases were filed in April and June 2024, respectively, and thus discovery should encompass documents through at least July 1, 2024. Microsoft's statement below (and brief) sidesteps the dispute identified by The Times's brief, which focuses on its refusal to produce documents about Microsoft's 2024 investments on the ground that 2024 is somehow too late. Third, and finally, The Times seeks an order requiring Microsoft to produce documents concerning the impact of Defendants' at-issue products on copyright owners. Microsoft has agreed to produce some documents addressing the specific three News Plaintiffs, which is unduly narrow. As a compromise, The Times has offered to narrow this request to the impact on the publishing industry, the news industry, cooking websites, and product recommendation websites, including the News Plaintiffs, but Micro

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							training and grounding data and negotiations related to them in Authors Guild v. OpenAI Inc., 23-cv-8292 (S.D.N.Y.), and Alter v. OpenAI Inc., 23-cv-10211, ECF 293 at 3 (S.D.N.Y.) (the "Class Cases"). With respect to RFP Nos. 113 and 114, in addition to being irrelevant and unnecessary because Microsoft has agreed to produce documents that may conceivably be relevant to the existence of a market, The Times's request for documents beyond unexecuted agreements and negotiations is not ripe for consideration.
							As to The Times's second category of documents, while The Times describes its request as merely asking for documents related to any ongoing investments by Microsoft in OpenAI, the specific request to which Microsoft objects is quite different: "Documents concerning whether Microsoft's potential or actual investments in OpenAI, including in the year[] 2024, would affect OpenAI's non-profit status, mission, or ability to engage in non-commercial activities." This request is, by its very terms, about OpenAI's state of mind, strategy, and "ability" and as such is best directed at OpenAI. Similarly, to the degree that some outside entity or regulator may have views on OpenAI's "non-profit status," it's hard to picture what documents Microsoft would have (beyond documents related to its investments in OpenAI, which are not at issue here because Microsoft has already agreed to produce such documents) that would be relevant.
							The Times's final request is not ripe for resolution by the Court. As The Times notes, Microsoft has agreed to produce any such documents that occur in the custodial documents that it collects and reviews based upon search terms, as well as to search for noncustodial documents regarding any impact on the News Plaintiffs. Additional meet and confer is required with respect to The Times's so-called compromise position so that the parties can try to reach agreement around questions such as how The Times is defining "the publishing industry, the news industry, cooking websites, and product recommendation websites."
				VI. Dispu	tes Related to Ne	ews Plaintiffs' Dis	covery Responses
383	1/13/25	<u>401</u>	OpenAI RFAs 36, 37, 39	OpenAI	The Times	OpenAI and The Times have submitted letter briefing.	OpenAI's position: OpenAI seeks an order compelling The Times to respond to RFAs aimed at whether it has licensed its asserted works to train any large language models. Whether the Times has licensed its works to train any LLMs is relevant to OpenAI's fair use defense because it goes directly to whether a workable market exists for

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							licensing data to train LLMs. In meet and confer, the only basis the Times offered for refusing to respond to these RFAs was that it did not understand the defined term "Generative AI"; tellingly, the Times has abandoned that justification. Now, the Times claims these RFAs are impermissible because they would require the Times to understand its own documents and whether or not it has been compensated for training data, and because there is some possibility the Times would have to amend its responses due to changed circumstances. Neither objection justifies the Times refusal to admit or deny these RFAs in accordance with FRCP 36. Plaintiff's position: OpenAI's motion should be denied for three reasons. First, the RFAs do not seek straightforward factual information; instead, responding to these requests would require The Times to interpret its content licenses, the terms of which are not always straightforward. Second, The Times has already agreed to produce its content licenses and related communications. Third, notwithstanding these objections, The Times has agreed to supplement these responses prior to the close of fact discovery. RFAs are not a discovery device, and responses to these RFAs are premature at best.
388	1/13/25	<u>410</u>	OpenAI RFPs 181 and 200	OpenAI	The Times	OpenAI and The Times have submitted letter briefing.	OpenAI's position: OpenAI seeks an order requiring The Times to produce two highly relevant categories of documents and communications. First, The Times should be required to produce documents relating to pricing changes for its products and the bases for those changes, which are relevant to The Times's demand for actual damages and OpenAI's fair use defense. Second, The Times should be ordered to produce documents and communications related to permitting search engines to access and index Times content. These documents will allow OpenAI to test The Times's allegation that search-engine copying is "different" from OpenAI's alleged conduct. The Times does not dispute that the documents it has already agreed to produce simply do not address the reasons for The Times's decision to permit search engines to access and copy Times content. Plaintiff's position: For the first category of documents (pricing documents), The Times has already agreed to produce documents showing the current pricing of its products and a long list of other damages-related discovery, including information about readership, online subscriptions,

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							online traffic, page views, and revenues. OpenAI's request for documents regarding anticipated pricing are irrelevant: anticipated pricing changes are not the basis for any damages claimed by The Times.
							The second category of documents (permissions given to search engines to copy Times content) is likewise irrelevant. Defendants' use of Times content to train their LLMs is entirely different from developing a search index that directs users to The Times's own website. Further, The Times's decision to allow some non-parties access to its content would have no bearing on The Times's refusal to provide such access to OpenAI. Nevertheless, The Times has already agreed to produce documents related to OpenAI's request, including documents about The Times's use of webmaster controls to prevent web crawlers from accessing or indexing The Times's website, analysis of the causes of increases, decreases, or fluctuations in click-through rates from search engines since 2019, documents sufficient to show the daily number of views within the nytimes.com domain resulting from click-throughs from search engines, and documents concerning the aggregate daily and monthly advertising revenue resulting from click-throughs from search engines.
394 – Sealed 397 – Public	1/13/25	413	OpenAI RFPs 182–183	OpenAI	The Times	OpenAI and The Times have submitted letter briefing.	OpenAI's position: OpenAI seeks an order compelling The Times to produce documents in its possession, custody or control from its business licensing consultant David B. Kogan. The Times relied heavily on Kogan to develop its strategy for licensing its published works, and his documents are directly relevant to the fair use factor four inquiry. The Times does not dispute that it has possession, custody, or control over Kogan's documents, which the Times has the legal right to obtain under its consulting contract with him. The Times's objections of burden and duplicativeness are unfounded and contradicted by the gaps in its production to date. Plaintiff's position: The Times has already agreed to produce responsive, non-privileged documents in its custody regarding the licensing of its asserted works for LLM training, including negotiations over such licenses. Burdening a third party with discovery for duplicative documents is disproportionate and unjustified.
							Nevertheless, The Times reached out to Mr. Kogan to inquire whether he has any documents regarding his consulting work for The Times

Case 1:24-cv-04872-SHS-OTW Document 199 Filed 01/17/25 Page 22 of 22

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							related to generative AI or other materials to which The Times is entitled under his consulting agreement. Mr. Kogan represented that he has no such documents. Accordingly, The Times has satisfied its discovery obligations with respect to OpenAI's requested documents.